Energy Minimization of Rigid-Geometry Polypeptides with Exactly Closed Disulfide Loops

K. D. GIBSON and H. A. SCHERAGA*

Baker Laboratory of Chemistry, Cornell University, Ithaca, NY 14853-1301

Received 1 April 1996; accepted 29 June 1996

ABSTRACT

A method has been developed for minimizing the energy of a polypeptide with rigid geometry while keeping all disulfide loops closed exactly. Exact closure of disulfide loops implies that some dihedral angles become implicit functions of the remaining dihedral angles in the polypeptide; the efficacy of the method is related to the manner in which the implicitly defined dihedral angles are chosen. The method has been used to find minimum-energy conformations of bovine pancreatic trypsin inhibitor, ribonuclease A, crambin, the defensin HNP3 dimer, and ω-conotoxin. For the first two proteins, the starting conformations for energy minimization had been derived previously from crystal structures using pseudopotentials to keep the disulfide loops almost closed. Starting conformations for the remaining three proteins were derived from their crystal or NMR structures by similar procedures. In all cases, the energy-minimized structures had a significantly and, in some cases, substantially, lower energy than the starting structures. The RMS deviations between the exactly closed energy-minimized structures and the crystal or NMR structures from which they were derived ranged from 0.9 Å to 1.9 Å, suggesting that the computed structures can serve as ''regularized'' native structures for these proteins. The energy of a ribonuclease derivative lacking the 65–72 disulfide bridge was minimized using the procedure; the result showed that this derivative has a low-energy structure with a conformation very close to that of native ribonuclease, and is consistent with its postulated role in the folding of ribonuclease. These results offer strong support for the validity of the rigid-geometry model in the studies of the conformational energy of proteins.

*Author to whom all correspondence should be addressed.
E-mail: has5@cornell.edu

Journal of Computational Chemistry, Vol. 18, No. 3, 403–415 (1997)

CCC 0192-8651 / 97 / 030403-13

Introduction

Molecular models of proteins for computational energy computations fall broadly into two categories: rigid-geometry models, in which bond lengths and bond angles are regarded as fixed and the only quantities that are allowed to vary are intramolecular dihedral angles and extramolecular rigid-body variables; and models with flexible geometry, in which the atoms or groups of atoms within each molecule are regarded as points in space connected by springlike forces. It has often been stated that the flexible models are the only correct ones, because bond lengths, and especially bond angles, in organic compounds are known to vary significantly from one compound to another; therefore, it is argued, the bond lengths and bond angles in any one molecule should show variations of the same order of magnitude when passing from one conformation to another. However, this argument ignores the enormous quantitative difference between the forces that determine molecular geometry and the forces that determine molecular conformation. The bond angles around any atom in a molecule are determined by the nature of the atoms that are joined to it by covalent bonds; if these atoms differ from one another in atomic number they will naturally influence the geometry around the central atom to a different extent. For example, the carbon atom in a methylene group within an alkane molecule is bonded to two carbon atoms and two hydrogen atoms, and it is hardly to be expected that the C—C—C, C—C—H, and H—C—H bond angles will all be equal; in fact, as is well known, the C—C—C angle is larger by 2–3° than the ideal tetrahedral angle expected for an sp^3 carbon atom bonded to four identical atoms.

The forces that determine the conformations of proteins and other molecules are in sharp quantitative contrast to those that determine molecular geometry. Peptide and protein conformations are determined by weak noncovalent forces such as van der Waals nonbonded interactions, hydrogen bonds, and long-range electrostatic forces between irregularly distributed charges, both within the peptide itself and between the peptide and its solvent.[1–3] These forces are at least one to two orders of magnitude weaker than the covalent forces that determine molecular geometry. The weak forces that determine protein conformation are also involved in crystal packing, and it is instructive to examine their effect in that situation. The crystal structure of C_6D_6 at temperatures down to 15 K has been determined by neutron diffraction, and the geometry of the (half-) molecule in the asymmetric unit is known rather precisely.[4] In the gas phase, benzene is a planar molecule with D_{6h} symmetry; all bond angles in the molecule are 120°. In the crystal, the largest deviation of any C—C—C bond angle from 120° is 0.25°; the C—C—D bond angles differ from 120° by 0.12° to 0.68°; two C atoms are 0.0039 Å from the plane of the other four; two D atoms are 0.0215 Å from the plane of the other four. The molecule in the asymmetric unit is subject to an essentially infinite number of noncovalent forces of van der Waals and electrostatic type, yet their effect on the geometry of the molecule is trivial in comparison with that of the covalent forces within the molecule itself.

The effect of noncovalent forces on peptide and protein geometry is exactly analogous. Weak noncovalent forces play a very minor role in determining the geometry of a peptide chain, although they play the major role in determining its conformation. The covalent geometry of amino acid residues varies from one amino acid to another, especially in the bond angles around the C^α atom,[5] but the geometry of each amino acid changes very little with conformation. There are isolated exceptions to this rule, which as the $C'—N—C^\alpha$ bond angle preceding a proline residue, which changes by 6° when the proline residue follows a *cis*-rather than a *trans*-peptide bond,[6] but these are not usually accompanied by similar changes in nearby portions of the peptide chain; thus, neither the C′—N bond length nor the bond angles around the C′ atom in the residue preceding the proline residue are affected, within experimental error, by the change in the conformation of the peptide bond. In computations with rigid-geometry models, allowance is easily made for variations in geometry between amino acid residues and for isolated changes in geometry. It follows that models which assume that the geometry of a protein is rigid are not only computationally convenient, as has been pointed out cogently by Swenson et al.[7] and by Palmer and Scheraga,[8] but also provide a satisfactory basis for conformational predictions based on energy. Such models also avoid the pitfalls inherent in some models with flexible geometry, in which bond angles in certain apparent low-energy equilibrium conformations of the force fields showed large deviations from their known experi-

mental values[9,10]; these deviations were sometimes large enough to suggest that the molecules in question should have significantly enhanced chemical reactivity.[11]

The rigid-geometry model for peptides and proteins that has been in use for two decades in this laboratory is that of ECEPP (Empirical Conformational Energy Program for Peptides).[5] In that model, the geometry of the peptide backbone and side chains was taken from high-resolution crystal structures of amino acids, peptides, and related compounds, with occasional revisions as the quality of the experimental data improved.[6,12] Associated with the rigid geometry of the model is a force field whose components consist of Lennard–Jones nonbonded interactions, electrostatic interactions between partial charges on atoms, special terms for hydrogen bonds, and intrinsic barriers to internal rotation.[5] The parameters of the force field have also been revised over the years.[6,12,13] Recently, another force field, which will be referred to as the NKS (No, Kang, Scheraga) potential, and which uses similar mathematical forms for its components but differs from the ECEPP potential in certain important respects, has been developed for use with the ECEPP model geometry.[14–18]

However, since its inception there has been an anomaly in the ECEPP force field; this relates to its manner of treating disulfide and other loops. In principle, if the geometry of a peptide can be regarded as rigid it should be possible to minimize its energy while keeping all bond lengths and bond angles unchanged, including any that are involved in disulfide loops. In the standard ECEPP force field, however, loop closure is achieved by the inclusion of a harmonic pseudopotential, or penalty function, to maintain certain interatomic distances as close as possible to their rigid-geometry values. For disulfide loops, the appropriate interatomic distances are those between the two S^{γ} atoms, and between each S^{γ} atom and the C^{β} atom of the opposite half-cystine residue (Fig. 1). The use of a pseudopotential not only introduces a spurious term into the energy without closing the loops exactly, but it calls into question the assumption underlying the use of rigid geometry by allowing certain bond lengths and bond angles to vary. Gō and Scheraga[19] showed how to close isolated loops exactly in peptides with rigid geometry, by solving a certain set of simultaneous nonlinear equations; for a given choice of the independent dihedral angles within the loop, their equations may have no solution (no loop closure)

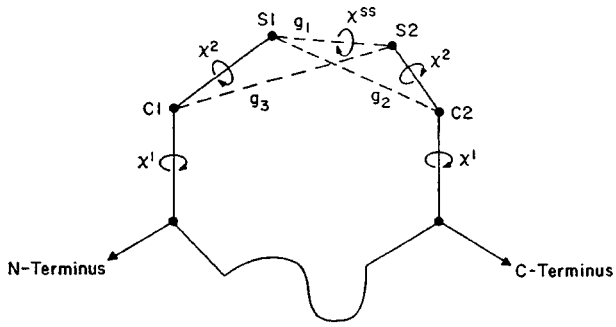

FIGURE 1. Identification of dihedral angles and of the atoms involved in the closure functions in a disulfide bridge.

or up to four different solutions. Palmer and Scheraga[8] showed that, whenever isolated loops can be closed, i.e., whenever the Gō–Scheraga equations have a solution, closure can be achieved computationally without changing bond angles, notwithstanding claims to the contrary.[20] However, neither study addressed the problem of minimizing the energy while keeping the loops closed.

In this article, we present a method for minimizing the energy of a rigid-geometry peptide with one or more disulfide loops, while keeping the disulfide loops exactly closed. In this way, the energy of the peptide is minimized without any bond length or bond angle being changed. Exact loop closure implies that some dihedral angles (six for each loop) must be treated as implicit functions of the remaining dihedral angles within the loop.[19] When the conditions for loop closure are defined in terms of the S^{γ}—S^{γ} and S^{γ}—C^{β} distances, as in Figure 1 and eqs. (1) (see Methods), three of the implicitly defined dihedral angles are determined automatically; these are the dihedral angles χ^2 of each half-cystine residue and the dihedral angle around the S—S bond (χ^{SS} in Fig. 1). The remaining three implicitly defined dihedral angles must be chosen from among the remaining dihedral angles within the loop.

The theoretical work of Gō and Scheraga,[19] and the extensive computational study by Dygert et al.[21] of the symmetrical closed-loop antibiotic, gramicidin S, both suggested that the dependent dihedral angles within a loop could be chosen arbitrarily. In our initial work with single disulfide loops from ribonuclease A or lysozyme, the dependent dihedral angles were taken to be the side-chain dihedral angles, χ^1, of the two half-cystine residues and the backbone dihedral angle, ϕ, of the C-terminal half-cystine residue; this choice was

completely satisfactory in all these tests. However, when the studies were extended to peptides with more than one loop we found that, except in the trivial case in which the loops were reducible in the sense of Benham and Jafri,[22] this choice of dependent dihedral angles failed even to close all the loops, let alone carry out energy minimization. The approach that finally succeeded consists of two steps. (1) To close all the loops exactly, Newton's method is used to locate the zero of a set of loop closure functions, three for each loop, with all backbone dihedral angles between the first and last cysteine residue, plus the side-chain dihedral angles, χ^1, of the cysteine residues, regarded as independent variables. (2) When all loops are closed exactly, a procedure developed by G. H. Golub, V. Klema, and G. W. Stewart in another context[23] is applied to pick out a subset of dihedral angles that will become implicit functions of the remaining dihedral angles during energy minimization.

Methods

CONDITIONS FOR LOOP CLOSURE

The constraints that must be satisfied when the mth disulfide loop is closed are taken to be:

$$\begin{aligned} g_1^{(m)}(\boldsymbol{\theta}) &= |\mathbf{r}_{S2}^{(m)} - \mathbf{r}_{S1}^{(m)}| - d_{SS} \\ g_2^{(m)}(\boldsymbol{\theta}) &= |\mathbf{r}_{S1}^{(m)} - \mathbf{r}_{C2}^{(m)}| - d_{SC} \\ g_3^{(m)}(\boldsymbol{\theta}) &= |\mathbf{r}_{S2}^{(m)} - \mathbf{r}_{C1}^{(m)}| - d_{SC} \end{aligned} \qquad (1)$$

In eqs. (1), $\mathbf{r}_i^{(m)}$ are cartesian coordinate vectors of atoms involved in the mth disulfide bridge, S1 and S2 refer to the S^γ atoms of the N-terminal and C-terminal half-cystine residues, and C1 and C2 refer to the C^β atoms of the N-terminal and C-terminal half-cystine residues, respectively (see Fig. 1). The quantities d_{SS} and d_{SC} are the rigid-geometry distances between the two S^γ atoms, and between an S^γ atom and the C^β atom on the opposite half-cystine residue, respectively; in the present study, these were set to their ECEPP values, namely $d_{SS} = 2.04$ Å and $d_{SC} = 3.052$ Å. If there are M disulfide loops there will be $3M$ similar equations of constraint. The vector $\boldsymbol{\theta}$ contains $3M$ dihedral angles, which are to be chosen from the N dihedral angles consisting of all backbone dihedral angles between the first and last half-cystine residues in the peptide chain, plus the side-chain dihedral angles, χ^1, of all the half-cystine residues. This method of formulating the conditions for loop closure is essentially the same as that used by Kostrowicki and Biernat[24] in their study of crown ethers. The functions $g_i^{(m)}$ also form the basis for the loop-closing pseudopotential used in the standard ECEPP algorithm.

The three constraints in eqs. (1) are equivalent to the six loop-closure conditions of Gō and Scheraga.[19] As noted in the Introduction, closure of each loop implies that six dihedral angles within it become implicit functions of the remaining dihedral angles, but when the closure conditions are formulated as in eqs. (1), three of the implicitly defined dihedral angles (both χ^2's and χ^{SS}) are determined automatically; i.e., once the positions of the C^α's, C^β's, and S^γ's are set, the dihedral angles, χ^2 and χ^{SS}, are defined uniquely. The remaining dihedral angles, three for each disulfide loop, make up the vector $\boldsymbol{\theta}$. To minimize the energy of the peptide while keeping all loops closed, it is necessary, first, to close all loops exactly, and then to choose the $3M$ implicitly defined dihedral angles in such a way that all disulfide loops remain closed during energy minimization. Our method for achieving these goals is described in the next section.

At the start of each iteration during energy minimization, the step to be taken in the space of independent dihedral angles will be tangent to the manifold of solutions to eqs. (1). However, a finite movement in the direction of this step may not remain on the manifold; hence, it will be necessary to change the implicitly defined dihedral angles so as to bring the conformation back on to the manifold. This is accomplished through the use of a modified Newton method to solve eqs. (1) starting from the values of those dihedral angles at the end of the previous iteration (for details, see below). When all disulfide loops are closed again, the gradient of the energy with respect to the independent dihedral angle variables can be calculated from the gradient with respect to all dihedral angles with the aid of the Jacobian matrix of eqs. (1), using a standard formula from the theory of implicit functions.[25] In this way, energy minimization is carried out in the space of rigid-body and independent dihedral angle variables only.

LOOP CLOSURE AND IMPLICITLY DEFINED DIHEDRAL ANGLES

We assume that a conformation has been found in which all disulfide loops are nearly closed, i.e., all constraints are satisfied within a reasonable tolerance. Such a conformation can be found with

the aid of a pseudo-potential to close the loops, as in the standard ECEPP potential. To close all disulfide loops exactly, we regard the $3M$ constraints as functions of all N backbone and side-chain dihedral angles. The equations to solve become:

$$\mathbf{g}(\boldsymbol{\eta}) = 0 \tag{2}$$

where:

$$\mathbf{g} = \left[g_1^{(1)}, g_2^{(1)}, g_3^{(1)}, g_1^{(2)}, g_2^{(2)}, g_3^{(2)}, \ldots, g_1^{(M)}, g_2^{(M)}, g_3^{(M)} \right]^T$$
$$\boldsymbol{\eta} = [\eta_1, \eta_2, \ldots, \eta_N] \tag{3}$$

and the η_i values refer to the N dihedral angles. Newton's method can be applied to solve eq. (2) if a solution exists for the equation:

$$\hat{\mathbf{J}}\mathbf{x} = -\mathbf{g}_c \tag{4}$$

where $\mathbf{g}_c$ is the current (nonzero) value of $\mathbf{g}$, $\mathbf{x}$ is a vector to be added to the current value of $\boldsymbol{\eta}$, and $\hat{\mathbf{J}}$ is the Jacobian matrix:

$$\hat{\mathbf{J}} = \begin{bmatrix} \frac{\partial g_1^{(1)}}{\partial \eta_1} & \frac{\partial g_1^{(1)}}{\partial \eta_2} & \cdots & \frac{\partial g_1^{(1)}}{\partial \eta_N} \\ \frac{\partial g_2^{(1)}}{\partial \eta_1} & \frac{\partial g_2^{(1)}}{\partial \eta_2} & \cdots & \frac{\partial g_2^{(1)}}{\partial \eta_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_3^{(M)}}{\partial \eta_1} & \frac{\partial g_3^{(M)}}{\partial \eta_2} & \cdots & \frac{\partial g_3^{(M)}}{\partial \eta_N} \end{bmatrix} \tag{5}$$

Since $\hat{\mathbf{J}}$ has $3M$ rows and N columns and $N \gg 3M$, eq. (4) will either have no solution or an infinite number of solutions. If there is no solution, the disulfide loops cannot be closed simultaneously, and the conformation is implausible, if not physically impossible. When eq. (4) can be solved for $\mathbf{x}$, it has a unique solution with minimum 2-norm, which can be found from the singular value decomposition (SVD) of $\hat{\mathbf{J}}$:

$$\hat{\mathbf{J}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \tag{6}$$

In eq. (6) $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\boldsymbol{\Sigma}$ is a $3M \times N$ matrix with zeros everywhere except on the leading diagonal. The SVD can also be written in terms of the columns, $\mathbf{u}_i$ of $\mathbf{U}$ and $\mathbf{v}_i$ of $\mathbf{V}$, and the singular values, σ_i of $\hat{\mathbf{J}}$:

$$\hat{\mathbf{J}} = \sum_{i=1}^{3M} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{7}$$

If the conformation is to maintain full flexibility when all loops are closed, eq. (4) must have a solution for any value of $\boldsymbol{\eta}$ within a neighborhood of the closed-loop conformation, and this implies that the rank of $\hat{\mathbf{J}}$ must be equal to $3M$. Thus, the singular values of $\hat{\mathbf{J}}$ are bounded away from zero, and the minimum 2-norm solution to eq. (4) is[26]:

$$\mathbf{x} = -\sum_{i=1}^{3M} \frac{\mathbf{u}_i \cdot \mathbf{g}_c}{\sigma_i} \mathbf{v}_i \tag{8}$$

We apply eq. (8) to find a search direction and step along it, iterating until all loops are closed; details are given in the next section. Because the search directions determined from eq. (8) have minimum 2-norm, we may expect the conformation with closed loops to approximate the starting conformation with open (but almost-closed) loops; if the latter has a low energy, the former should also have a relatively low energy.

Once the loops have been closed, it is necessary to choose the set of $3M$ dihedral angles θ_i which are to become implicit functions of the remaining N—$3M$ dihedral angles. For this to be possible, the Jacobian matrix, $\mathbf{J}$, must be nonsingular, where:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial g_1^{(1)}}{\partial \theta_1} & \frac{\partial g_1^{(1)}}{\partial \theta_2} & \cdots & \frac{\partial g_1^{(1)}}{\partial \theta_{3M}} \\ \frac{\partial g_2^{(1)}}{\partial \theta_1} & \frac{\partial g_2^{(1)}}{\partial \theta_2} & \cdots & \frac{\partial g_2^{(1)}}{\partial \theta_{3M}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_3^{(M)}}{\partial \theta_1} & \frac{\partial g_3^{(M)}}{\partial \theta_2} & \cdots & \frac{\partial g_3^{(M)}}{\partial \theta_{3M}} \end{bmatrix} \tag{9}$$

Furthermore, $\mathbf{J}$ should remain nonsingular throughout energy minimization even though the values of θ_i change. If O denotes the set of singular $3M \times 3M$ matrices, we wish to ensure that the initial $\mathbf{J}$ is bounded well away from O, so as to reduce the probability that $\mathbf{J}$ becomes singular during the course of energy minimization. The distance (with respect to the 2-norm) of $\mathbf{J}$ from O is equal to $\sigma_{3M}(\mathbf{J})$, the smallest singular value of $\mathbf{J}$. Thus, the criterion for choosing the implicitly defined dihedral angles is that the singular values of the resulting Jacobian matrix be bounded well away from zero.

To solve the problem of choosing the implicit dihedral angles using this criterion, we apply a procedure developed by Golub, Klema, and

Stewart (see Golub and Van Loan[23]) in connection with the rank-deficient linear least-squares problem, as mentioned in the Introduction. Let $\mathbf{P}$ be an $N \times N$ permutation matrix, and let $\hat{\mathbf{J}}\mathbf{P}$ be partitioned into:

$$\hat{\mathbf{J}}\mathbf{P} = [\mathbf{J}, \mathbf{B}] \tag{10}$$

If $\hat{\mathbf{V}} = \mathbf{P}^T\mathbf{V}$, where $\mathbf{V}$ is the orthogonal $N \times N$ matrix on the right hand side of eq. (6), and if $\hat{\mathbf{V}}$ is partitioned into:

$$\hat{\mathbf{V}} = \begin{bmatrix} \hat{\mathbf{V}}_{11} & \hat{\mathbf{V}}_{12} \\ \hat{\mathbf{V}}_{21} & \hat{\mathbf{V}}_{22} \end{bmatrix} \tag{11}$$

where $\hat{\mathbf{V}}_{11}$ is a $3M \times 3M$ submatrix, then[27]:

$$\sigma_{3M}(\mathbf{J}) \geq \frac{\sigma_{3M}(\hat{\mathbf{J}})}{\|\hat{\mathbf{V}}_{11}^{-1}\|_2} \tag{12}$$

(since $\hat{\mathbf{V}}$ is orthogonal and $\hat{\mathbf{V}}_{11}$ is a square submatrix of $\hat{\mathbf{V}}$, $\|\hat{\mathbf{V}}_{11}\|_2 \leq 1$ and $\|\hat{\mathbf{V}}_{11}^{-1}\|_2 \geq 1$, regardless of how $\hat{\mathbf{V}}_{11}$ is chosen). The original problem is now transformed into the problem of finding a permutation matrix $\mathbf{P}$ that makes $\|\hat{\mathbf{V}}_{11}^{-1}\|_2$ as small as possible; this is equivalent to requiring that $\hat{\mathbf{V}}_{11}$ be well-conditioned. Golub, Klema, and Stewart suggested the following heuristic solution: partition $\mathbf{V}$ from eq. (6) into:

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \tag{13}$$

where $\mathbf{V}_{11}$ is a $3M \times 3M$ submatrix, set:

$$\mathbf{W} = \begin{bmatrix} \mathbf{V}_{11}^T & \mathbf{V}_{21}^T \end{bmatrix} \tag{14}$$

and compute the QR with column pivoting factorization[28] of $\mathbf{W}$:

$$\mathbf{W}\mathbf{P} = \mathbf{Q}[\mathbf{R}_{11}, \mathbf{R}_{12}] \tag{15}$$

In eq. (15), $\mathbf{Q}$ is a $3M \times 3M$ orthogonal matrix and $\mathbf{R}_{11}$ is a $3M \times 3M$ nonsingular upper triangular matrix. The heuristic value of this factorization is that $\|\hat{\mathbf{V}}_{11}^{-1}\|_2 = \|\mathbf{R}_{11}^{-1}\|_2$ and the factorization almost always produces a well-conditioned matrix $\mathbf{R}_{11}$; hence, from eq. (12), $\mathbf{J}$ will be bounded well away from the set O of singular matrices.

In sum, the algorithm for closing all loops and choosing a set of dihedral angles that will become implicit functions is: (1) compute the SVD of $\hat{\mathbf{J}}$ and use it to define a Newton search direction, iterating until all loops are closed; (2) partition the matrix $\mathbf{V}$ as indicated in eq. (13), form the matrix $\mathbf{W}$ from this partition, and compute its QR with column pivoting factorization; (3) use the first $3M$ column permutations that arise from the factorization to define the dihedral angles that will become implicit functions during energy minimization.

ALGORITHMS FOR LOOP CLOSURE

The initial procedure for closing all loops uses line searches with backtracking, the directions of search being defined by eq. (8). The criterion for accepting each step is that the sum-of-squares function, $(1/2)\mathbf{g}^T\mathbf{g}$, shows a sufficient decrease; if not, backtracking is carried out according to a standard procedure.[29] The line searches are iterated until all loops are closed. To maintain or restore loop closure during energy minimization, a more elaborate procedure was found preferable. In this procedure, searches in the Newton direction are combined with minimization of the sum-of-squares function, with a trust region to limit the extent of each line search and a dogleg strategy to choose the direction of search. Full details are given in Dennis and Schnabel.[29] In all cases, the criterion for accepting a closed-loop conformation was that every constraint equation should be satisfied within 1.12×10^{-8} Å.

COMPUTATIONAL DETAILS

Energy minimizations were carried out by using the SUMSL (secant unconstrained minimization solver) routine of Gay[30]; as noted above, the set of variables for these minimizations consisted of all rigid body variables and all independent dihedral angles. Comparisons of CPU time showed that the closed loop algorithm and the standard ECEPP algorithm consumed the same amount of CPU time, within 10%, in all cases examined. Structures were superposed using an algorithm based on the SVD.[9,31]

Results

The algorithm has been applied to five proteins of known structure: bovine pancreatic trypsin inhibitor (BPTI); bovine pancreatic ribonuclease A (RNase A), crambin; defensin HNP-3; and ω-conotoxin. The first three and the last protein are monomers; the fourth protein is a dimer in the crystal and probably in solution.[32] The arrangement of disulfide loops in these proteins is sketched

in Figure 2. Energy minimization with BPTI and RNase A were started from conformations that had been computed previously in this laboratory using the standard ECEPP algorithm with a loop-closing potential; therefore, energy minimization with the closed-loop algorithm was carried out with both the ECEPP/3 potential and the NKS potential. For the other three proteins, only the NKS potential was employed.

BOVINE PANCREATIC TRYPSIN INHIBITOR

The starting conformation had been derived earlier from the 4PTI crystal structure[33] by Meirovitch and Scheraga[34] and Ripoll et al.[35]; the energy of this conformation had been minimized with the standard ECEPP/3 algorithm. In this conformation, the absolute values of the closure functions [see eqs. (1) under Methods] ranged from 0.001 Å to 0.020 Å. When the ECEPP/3 energy of this conformation was minimized with the closed-loop algorithm, it decreased by a further 14 kcal/mol; a similar run with the NKS potential led to a decrease of 126 kcal/mol (Table I). The RMS deviations between all heavy atoms in the initial and final conformations were about 0.9 Å, and the two energy-minimized conformations differed from each other by a similar amount (Table II). The dihedral angles that were chosen to be implicit functions were Pro13 ψ, Lys15 ϕ, Arg17 ψ, Lys26 ψ, Ala27 ψ, Gly36 ψ, Cys38 ω, Asp50 ϕ, and Arg53 ϕ. The positions of these dihedral angles in one of the final conformations are shown in Figure 3; the arrows in Figure 2A point to the residues that contain them.

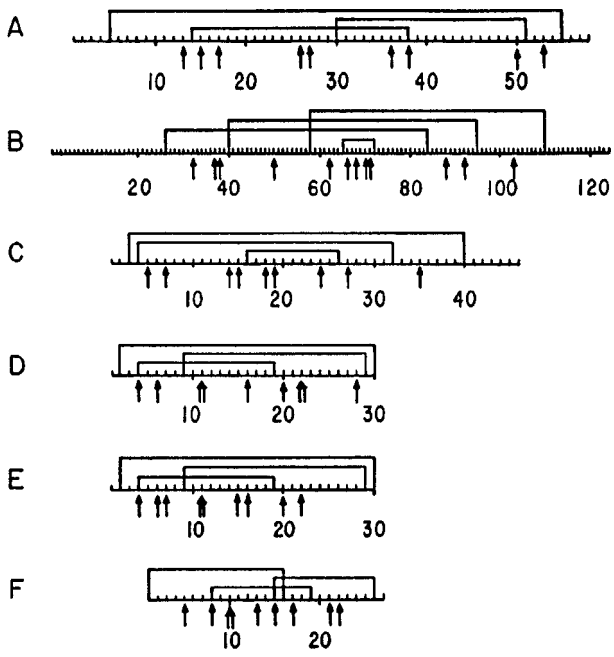

FIGURE 2. Schematic representation of disulfide loops in five proteins. (A) BPTI; (B) RNase A; (C) crambin; (D) and (E) defensin strands A and B; (F) ω-conotoxin. Residue numbers of defensin strands start at 2, to be consistent with homologous defensins in other mammals.[32] Arrows indicate residues in which one or more dihedral angles were chosen to be implicit functions of the remaining dihedral angles (see Methods).

BOVINE PANCREATIC RIBONUCLEASE A

The initial conformation for this study had been used by Talluri (private communication) to test the procedure used by Talluri et al.[36] for characterizing the structure of a three-disulfide intermediate in the folding pathway of ribonuclease. The conformation was obtained by measuring a set of NOE distance constraints from the 7RSA crystal structure of bovine ribonuclease A,[37] and deducing a structure from these by applying the program DIANA[38] followed by restrained energy minimization.[39] The RMS deviation of this structure from the crystal structure was 0.52 Å (see Table V). The absolute values of the loop closure functions ranged from 0.001 Å to 0.071 Å. The peptide sequence used in deriving this conformation and in our subsequent energy minimizations differed

TABLE I.
Energies of Initial and Final Conformations of BPTI and RNase A.

Protein	Potential	Conformation[a]	Energy (kcal / mol)
BPTI	ECEPP / 3	Initial, open	−513.0[b]
		Initial, closed	−511.2
		Final	−526.8
	NKS	Initial, open	−1187.7[b]
		Initial, closed	−1172.3
		Final	−1313.5
RNase A	ECEPP / 3	Initial, open	−972.5[b]
		Initial, closed	−962.3
		Final	−1044.0
	NKS	Initial, open	−1430.1[b]
		Initial, closed	−1407.8
		Final	−1806.7

[a]Initial, open: conformation before loop closure; initial, closed: conformation after loop closure; final: after energy minimization with exact loop closure.
[b]Excluding contribution from loop-closure pseudopotential.

TABLE II.
Comparisons of BPTI and RNase A Conformations.

Protein	Conformation no. 1	Conformation no. 2	RMSD[a] (Å)
BPTI	Initial	ECEPP/3 − minimized	0.929
	Initial	NKS-minimized	1.289
	ECEPP/3 − minimized	NKS-minimized	1.371
RNase A	Initial	ECEPP/3 − minimized	0.903
	Initial	NKS-minimized	1.607
	ECEPP/3 − minimized	NKS-minimized	1.308

[a]RMS deviation, all heavy atoms.

from that of bovine ribonuclease A through the substitutions Met13 → Gln, Met29 → Gln, and Met30 → Gln.

As with BPTI, energy minimization with analytically closed loops was carried out separately with the ECEPP/3 potential and the NKS potential. Since the starting conformation did not correspond to a minimum of the standard ECEPP energy, the decrease in energy as a result of energy minimization was substantial in both cases, amounting to 72 kcal/mol with the ECEPP potential and 376 kcal/mol with the NKS potential (Table I). As with BPTI, the RMS deviations between all heavy atoms in the initial and final conformations were close to 1.0 Å with either potential, and the RMS deviation between the heavy atoms in the two energy-minimized conformations was also close to 1.0 Å (Table II). The dihedral angles that were chosen to be implicit functions were Ser32 ψ, Lys37 ω, Asp38 ψ, Ser50 ϕ, Asn62 ϕ, Lys66 ω, Gly68 ψ, Thr70 ϕ, Asn71 ω, Glu88 ω, Ser92 ϕ, and Asn103 ψ. The positions of these dihedral angles in one of the final conformations are shown in Figure 4, and their places in the amino acid sequence are indicated in Figure 2B.

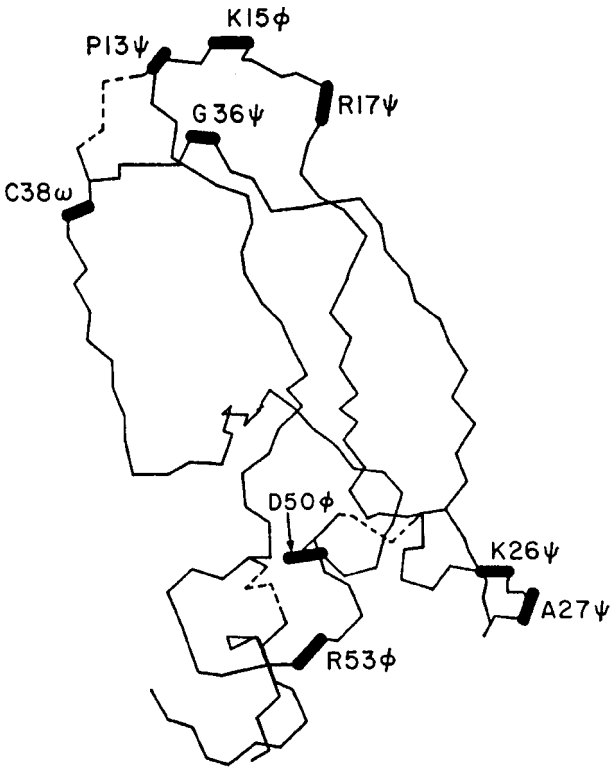

FIGURE 3. NKS energy-minimized conformation of backbone skeleton and disulfide bridges of BPTI, showing dependent dihedral angles. The dashed lines indicate the disulfide bridges.

CRAMBIN

For this and the next two peptides, no good starting conformation was readily available. A conformation of crambin to which the closed loop algorithm could be applied was obtained by the following procedure. (1) A set of dihedral angles was deduced from the heavy atom coordinates of the 1CRN crystal structure,[40,41] and all dihedral angles (those involving unseen hydrogen atoms) for which values could not be found in this way were set to 180°. (2) Starting from these dihedral angles, the NKS energy of fully reduced crambin (with cysteine replacing each half-cystine residue)

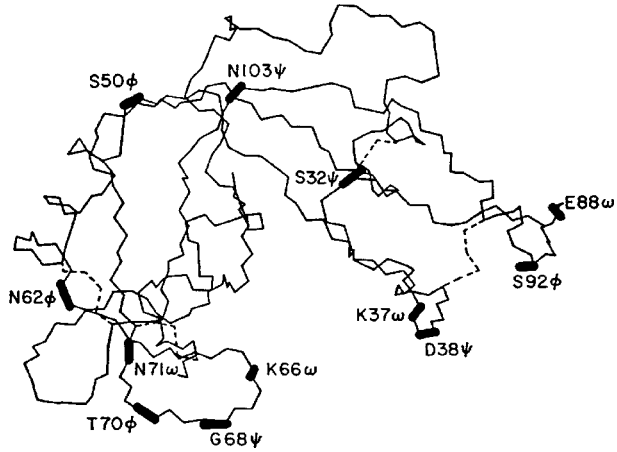

FIGURE 4. NKS energy-minimized conformation of backbone skeleton and disulfide bridges of RNase A, showing dependent dihedral angles. The dashed lines indicate the disulfide bridges.

was minimized. (3) Starting from the final conformation reached in step 2, the NKS energy of (oxidized) crambin was minimized with a pseudopotential to close disulfide loops, as in the standard ECEPP potential.

The final conformations from steps (2) and (3) were stationary points but probably not true local minima; nevertheless, the final conformation from step (3) served as a satisfactory starting conformation for the closed-loop algorithm, since the absolute values of the loop closure functions ranged from 0.005 Å to 0.119 Å. Table III shows the energies of the conformations from steps (2) and (3), and of the final conformation (which was a true local minimum). The dihedral angles that were chosen to be implicit functions were Pro5 ψ, Ile7 ω, Asn14 ω, Val15 ψ, Leu18 ψ, Pro19 ψ, Ala24 ψ, Ala27 ψ, and Ile35 ψ; their positions in the final conformation and in the sequence are shown in Figures 5 and 2C, respectively.

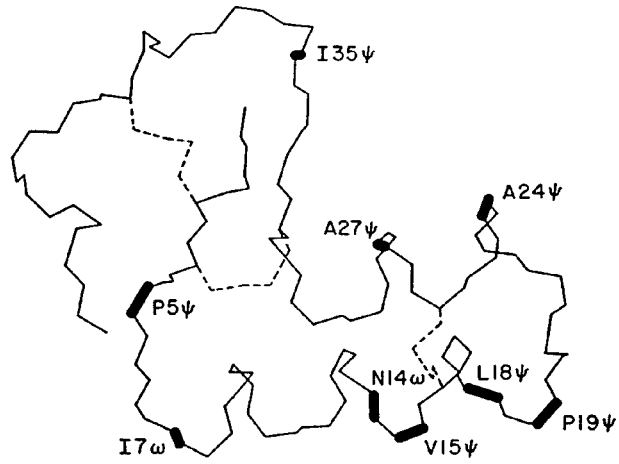

FIGURE 5. Energy-minimized conformation of backbone skeleton and disulfide bridges of crambin, showing dependent dihedral angles. The dashed lines indicate the disulfide bridges.

DEFENSIN

The method used to obtain a suitable starting conformation for crambin failed when applied to defensin (i.e., a starting conformation, which allowed the loops to be closed, could not be found), perhaps because the relatively low resolution of the 1DFN crystal structure[32] did not lead to a good set of initial dihedral angles. However, the following procedure led to a satisfactory starting conformation. (1) A set of dihedral angles, and rigid-body variables for strand B relative to strand A, was deduced from the crystal structure where possible, other dihedral angles being set to 180° (including the side-chain dihedral angles of Arg15 which was not seen in either strand; the residue numbering is that of Hill et al.[32]). (2) A weighted sum of squares of the distances between the heavy atoms in the computer-generated structure and the crystal structure was minimized. (3) A further restrained energy minimization of the NKS potential was carried out, with a pseudopotential to close disulfide loops. In this step, the least-squares function was the restraint, and the relative weights of this function and the energy were 0.99 and 0.01, respectively.

The absolute values of the closure functions at the end of the third step ranged from 0.001 Å to 0.041 Å. Table IV shows the energies of this structure and of the final structure obtained with the closed-loop algorithm. The two molecules of water that form bridges between the two strands in the crystal were not included in the computations because of the lack of a satisfactory potential for H_2O. The defensin dimer in the crystal has a twofold axis which is not crystallographic, and this was almost preserved in the final energy-minimized structure. The RMS deviation between all backbone C, C^{α}, and C′ atoms of the two strands in the final structure was 0.683 Å, as compared with 0.806 Å in the crystal structure. Superposition of the backbone atoms of strand B on to those of strand A in the crystal resulted in a rotation angle of 179.4° and a translation vector of length 26.1 Å; the corresponding numbers for the computer-generated structure were 174.4° and 27.3 Å, respectively. The slight difference in conformation between the two strands was reflected in the choice of dihedral angles that became implicit functions.

TABLE III.
Conformational Energies of Crambin.

Protein	Loop Closure	Energy (kcal / mol)
Reduced crambin	—	−158.5 (step 2)
Crambin	Pseudopotential	−380.3[a] (step 3)
Crambin	Exact, after loop closure[b]	−365.2
Crambin	Exact, energy-minimized[c]	−519.8

[a]Excluding contribution from loop-closure pseudopotential.
[b]After loop closure but before energy minimization.
[c]After both loop closure and energy minimization.

TABLE IV.
Conformational Energies of Defensin and ω-Conotoxin.

Protein	Conformation[a]	Energy (kcal / mol) Strand A	Strand B	Interstrand	Total
Defensin	Initial, open	−735.8	−738.4	−78.2	−1552.4[b]
	Initial, closed	−734.1	−736.4	−77.6	−1548.1
	Final	−749.4	−761.6	−84.7	−1595.7
ω-Conotoxin	Initial, open				−211.8[b]
	Initial, closed				−151.2
	Final				−248.9

[a]Initial, open: conformation before loop closure (a pseudopotential was used to generate the energy-minimized "initial, open" conformation); initial, closed: conformation after loop closure; final: after energy minimization with exact loop closure.
[b]Excluding contribution from loop-closure pseudopotential.

In both strands, Cys5 ϕ, Ile7 ψ, Ala12 ψ and ω, Ile21 ψ and Gln23 ψ were chosen; however, in strand A, the remaining implicitly defined dihedral angles were Tyr17 ω, Gln23 ω, and Phe29 ω, whereas in strand B they were Pro8 ω,Arg16 ω, and Gly24 ω. The positions of these dihedral angles in the final conformation are shown in Figure 6; the residues in which they are located are shown in Figures 2D and E.

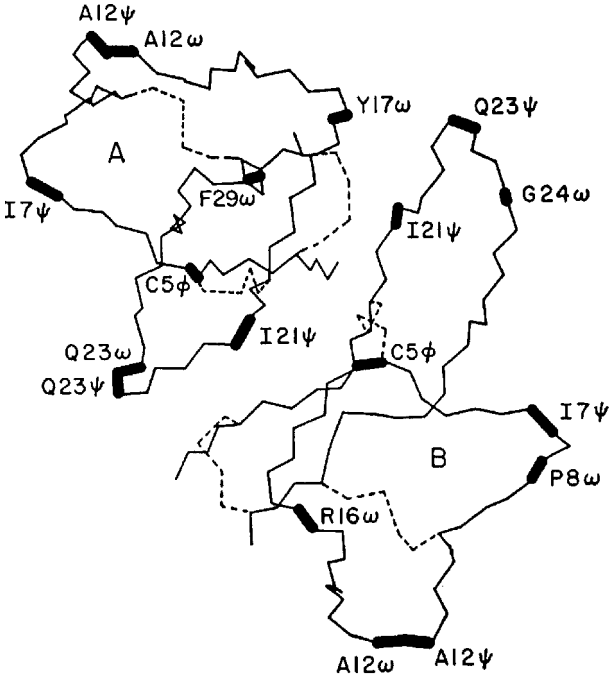

FIGURE 6. Energy-minimized structure of backbone skeleton and disulfide bridges of defensin dimer, showing dependent dihedral angles. The numbering system is that of Hill et al.[32] The dashed lines indicate the disulfide bridges.

ω-CONOTOXIN

A suitable starting conformation for energy minimization of ω-conotoxin was obtained in a manner similar to that used for defensin. Dihedral angles were calculated from the first NMR structure of Pallaghy et al.,[42] and used as input for weighted least-squares minimization, followed by restrained NKS energy minimization with a pseudopotential to close disulfide loops. The absolute values of the closure functions at the end of this step ranged from 0.019 Å to 0.103 Å. Energy minimization with the closed loop algorithm led to a decrease in energy of 37 kcal/mol compared to the conformation whose energy had been minimized with the pseudopotential (Table IV). The dihedral angles chosen to become implicit functions were Gly5 ψ, Cys8 ω, Hyp10 ψ and ω, Tyr13 ω, Cys15 ϕ, Arg17 ψ, Hyp21 ω, and Tyr22 ψ; their positions in the final conformation are shown in Figure 7, and in the amino acid sequence in Figure 2F.

COMPARISON WITH CRYSTAL OR NMR STRUCTURES

The backbone atoms of several of the computed structures were superposed on to the observed structures from which they were derived; RMS deviations are presented in Table V.

The conformation of BPTI that was used as the starting point for energy minimization in the present work has been used as a "regularized" native structure in several studies from this laboratory.[35,43,44] The RMS deviation of its backbone atoms from the corresponding atoms in the 4PTI structure is 1.75 Å. The two energy-minimized conformations that were obtained with the

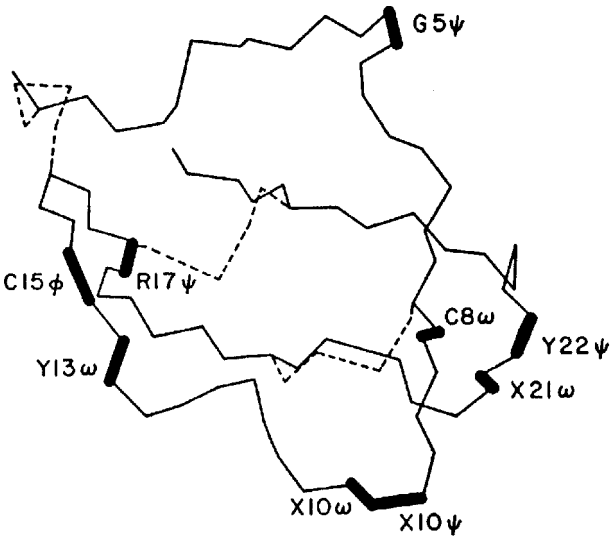

FIGURE 7. Energy-minimized conformation of backbone skeleton and disulfide bridges of ω-conotoxin, showing dependent dihedral angles. Residues denoted by **X** are hydroxyproline. The dashed lines indicate the disulfide bridges.

closed-loop algorithm had RMS deviations within 0.2 Å of this value. Thus, energy minimization with exactly closed loops lowered the conformational energy without greatly affecting the resemblance of the computed structures to the crystal structure.

Unlike the initial conformation of BPTI, the starting conformation for energy minimization of RNase A did not correspond to a local minimum of any energy potential because of the imposition of NOE constraints. By design, its RMS deviation from the 7RSA crystal structure of RNase A was quite small. Energy minimization with the closed loop algorithm, with either the ECEPP/3 or NKS potential, led to a final conformation whose RMS deviation from the crystal structure was close to 1.0 Å. This deviation is about double that of the starting conformation, but well within the range of values that are generally observed in structural comparisons of this type.

For each of the remaining three proteins, crambin, defensin and conotoxin, the starting conformation for application of the closed-loop algorithm was not a minimum of an energy function, and the opportunity for appreciable conformational change was always present. Nevertheless, in all three cases, the RMS deviation between the energy-minimized conformations obtained with the closed-loop algorithm and the crystal or NMR structures from which they were derived were in line with those observed for BPTI and ribonuclease, being 1.65 Å for crambin, 0.87 Å for defensin, and 1.14 Å for conotoxin.

We conclude that the energy-minimized conformation obtained with the closed-loop algorithm for any of these five proteins could serve as a satisfactory ''regularized'' native structure for use in further study.

TABLE V.
Comparison of Computed and Experimental Structures.

Protein	Experimental structure[a]	Conformation	RMSD (Å)[b]
BPTI	4PTI	Initial	1.754
		ECEPP/3 − minimized	1.897
		NKS-minimized	1.536
RNase A	7RSA	Initial	0.522
		ECEPP/3 − minimized	0.923
		NKS-minimized	1.283
Crambin	1CRN	NKS-minimized	1.654
Defensin	1DFN	NKS-minimized	0.873
ω-Conotoxin	1CCO[c]	NKS-minimized	1.138

[a] Protein Data Bank identifier.
[b] All backbone N, C^{α}, and C′ atoms.
[c] NMR structure no. 1.

THREE-DISULFIDE RIBONUCLEASE A

As a final example, we present the results of NKS energy minimization of a ribonuclease molecule in which Cys65 and Cys72 have been replaced by cysteine residues. The three-disulfide species lacking the 65–72 disulfide bridge (des-[65–72]-RNase) is one of the major three-disulfide intermediates in the refolding pathway of ribonuclease.[45,46] A ribonuclease derivative in which Cys65 and Cys72 were replaced by blocked cysteine residues has been shown from NMR data to have a conformation resembling that of native ribonuclease.[36] It seemed of interest to minimize the energy of a ribonuclease molecule lacking the 65–72 disulfide bridge to determine whether its absence would lead to a significant conformational change.

The starting conformation for application of the closed loop algorithm was the same as that for RNase A itself, except that the values of χ^2 for cysteines 65 and 72 were chosen so as to direct the HS^{γ} atoms away from each other. Even so, the

initial energy was large and positive (1471 kcal/mol), presumably because of atomic overlap between the two S^{γ} atoms. The final NKS energy was −1812.0 kcal/mol, which differs from that of the energy-minimized conformation of ribonuclease (Table I) by 5.3 kcal/mol. The final conformations of ribonuclease and des-[65–72]-RNase were compared by superposition. The RMS deviation of all heavy atoms in residues 1–64 and 73–124 inclusive was 1.545 Å between the two structures; the RMS deviation of all heavy atoms within the 65–72 loop was 0.439 Å. This result supports the view that des-[65–72]-RNase has a low-energy conformation similar to the native conformation of RNase A itself, and is consistent with the proposed folding pathway in which des-[65–72]-RNase is one of the immediate precursors of the fully oxidized native molecule.[45,46]

Discussion

The most important aspect of the work presented here is that it demonstrates that the model of protein structure in which all bond lengths and bond angles are fixed is computationally feasible and practical, even when the protein has several closed loops whose conformations are interdependent. This adds strong support to the view that rigid-geometry models, in which the geometry is matched to experimental data, can be used to predict the conformation of any peptide or protein. The work reported here deals only with disulfide loops, but there is clearly no obstacle to extending the algorithm to cover any type of intramolecular closed loop.

There are several features of interest about the dihedral angles that were chosen by the algorithm to become implicit functions of the remaining dihedral angles. First, out of 57 such dihedral angles in BPTI, RNase A, crambin, defensin, and conotoxin, only five were in half-cystine residues: one in BPTI; two in conotoxin; and one in each strand of defensin. Furthermore, all of these dihedral angles were in the backbone rather than the side chain. It is also noteworthy that only one of the dependent dihedral angles in half-cystine residues could have any direct effect on the disulfide bridge involving that half-cystine residue; this is Cys8 ω in conotoxin, which lies within the 8–19 loop of that protein. Another interesting feature is the tendency of the implicitly defined dihedral angles to occur in imino acid residues, in spite of the reduced flexibility of these residues in the rigid-geometry model. Thus, implicitly defined dihedral angles were found in one of the four proline residues in BPTI, in two of the five proline residues in crambin, in the only proline residue in defensin B, and in two of the three hydroxyproline residues in conotoxin.

However, the most significant feature of the choice of implicitly defined dihedral angles is their strong tendency to occur in surface-exposed residues. This is true of nearly all the dihedral angles in BPTI (Fig. 3), and RNase A (Fig. 4), and of most of those in crambin (Fig. 5). There is also a tendency to occur in exposed loops; for example, in residues 26 and 27 of BPTI; 18, 19, and 24 of crambin; 12, 21, and 23 of defensin; 5, 10, 21, and 22 of conotoxin; and at each end of the RNase A molecule. A further tendency to occur in clusters is also evident; for example, two distinct clusters at opposite ends of the BPTI molecule and the RNase A molecule, one major cluster at one end of the crambin molecule, and two clusters in the conotoxin molecule. Only in defensin are the implicitly defined dihedral angles distributed fairly evenly through the molecule. In BPTI and RNase A, the clustering closely parallels the distribution of disulfide bridges within the molecule, but this does not seem to be true for crambin. The tendency of the implicitly defined dihedral angles to occur in surface residues, clusters, and loops is rather remarkable considering that the procedure which chose them was entirely mathematical in nature, and did not involve the energy function or any other physicochemical property of the molecules in any way.

Finally, it should be noted that the choice of implicitly defined dihedral angles depends on the starting conformation. Even a small difference in the conformational changes leading to exact loop closure can have an effect on this choice. For example, in the computation with des-[65–72]-RNase, eight of the nine dihedral angles that were chosen to become defined implicitly were identical with eight of the implicitly defined dihedral angles in the computations with RNase A. The ninth dihedral angle was Glu60 ψ, which was chosen instead of Asn62 ϕ. The absence of the 65–72 disulfide loop led to a small change in the path taken by the loop closure procedure, and this in turn led to an alternative choice for one of the dependent dihedral angles. The three dihedral angles that were implicitly defined in RNase A, but not in des-[65–72]-RNase, were Gly68 ψ, Thr70 ϕ, and Asn71 ω; as expected, these are all inside the 65–72 disulfide loop.

Acknowledgments

This work† was supported by NIH Grant No. GM-14312, and by NSF Grant No. MCB95-13167. The computations were carried out on an IBM SP2 supercomputer of the Center for Theory and Simulation in Science and Engineering at Cornell University, which is funded in part by the National Science Foundation, the State of New York, the IBM Corporation, and members of its Corporate Research Institute, with additional funding from the National Institutes of Health.

References

1. G. Némethy and H. A. Scheraga, *Q. Rev. Biophys.*, **10**, 239 (1977).
2. M. Levitt, *Annu. Rev. Biophys. Bioeng.*, **11**, 251 (1982).
3. C. L. Brooks, III, M. Karplus, and B. M. Pettitt, *Adv. Chem. Phys.*, **71**, 1 (1988).
4. G. A. Jeffrey, J. R. Ruble, R. K. McMullan, and J. A. Pople, *Proc. R. Soc.* (London) **A414**, 47 (1987).
5. F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
6. G. Némethy, K. D. Gibson, K. A. Palmer, C. N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, and H. A. Scheraga, *J. Phys. Chem.*, **96**, 6472 (1992).
7. M. K. Swenson, A. W. Burgess, and H. A. Scheraga, In *Frontiers in Physicochemical Biology*, B. Pullman, Ed., Academic Press, New York, 1978, p. 115.
8. K. A. Palmer and H. A. Scheraga, *J. Comput. Chem.*, **12**, 505 (1991).
9. I. K. Roterman, K. D. Gibson, and H. A. Scheraga, *J. Biomolec. Struct. Dyn.*, **7**, 391 (1989).
10. I. K. Roterman, M. H. Lambert, K. D. Gibson, and H. A. Scheraga, *J. Biomolec. Struct. Dyn.*, **7**, 421 (1989).
11. K. D. Gibson and H. A. Scheraga, *J. Biomol. Struct. Dyn.*, **8**, 1109 (1991).
12. G. Némethy, M. S. Pottle, and H. A. Scheraga, *J. Phys. Chem.*, **87**, 1883 (1983).
13. M. J. Sippl, G. Némethy, and H. A. Scheraga, *J. Phys. Chem.*, **88**, 6231 (1984).
14. K. T. No, J. A. Grant, and H. A. Scheraga, *J. Phys. Chem.*, **94**, 4732 (1990).
15. K. T. No, J. A. Grant, M. S. Jhon, and H. A. Scheraga, *J. Phys. Chem.*, **94**, 4740 (1990).
16. K. T. No, O. Y. Kwon, S. Y. Kim, M. S. Jhon, and H. A. Scheraga, *J. Phys. Chem.*, **99**, 3478 (1995).

†The computer program may be obtained from QCPE. Contact: QCPE, Creative Arts Bldg. 181, Indiana University, Bloomington, IN 47405; E-mail: qcpe@indiana.edu and request program QCPE 664 (1996) by K. D. Gibson and H. A. Scheraga.

17. K. T. No, O. Y. Kwon, S. Y. Kim, K. H. Cho, C. N. Yoon, Y. K. Kang, K. D. Gibson, M. S. Jhon, and H. A. Scheraga, *J. Phys. Chem.*, **99**, 13019 (1995).
18. Y. K. Kang, K. T. No, and H. A. Scheraga, *J. Phys. Chem.* (in press).
19. N. Gō and H. A. Scheraga, *Macromolecules*, **3**, 178 (1970).
20. R. E. Bruccoleri and M. Karplus, *Macromolecules*, **18**, 2767 (1985).
21. M. Dygert, N. Gō, and H. A. Scheraga, *Macromolecules*, **8**, 750 (1975).
22. C. J. Benham and M. S. Jafri, *Prot. Sci.*, **2**, 41 (1993).
23. G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1985, p. 414.
24. J. Kostrowicki and J. F. Biernat, *J. Incl. Phenom. Molec. Recogn. Chem.*, **11**, 205 (1991).
25. M. Spivak, *Calculus on Manifolds*, Benjamin, New York, 1965, pp. 40–43.
26. Reference 23, p. 187.
27. Reference 23, p. 416.
28. Reference 23, p. 164.
29. J. E. Dennis Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983, chaps. 5–7.
30. D. M. Gay, *ACM Trans. Math. Software*, **9**, 503 (1983).
31. Reference 23, pp. 425–426.
32. C. P. Hill, J. Yee, M. E. Selsted, and D. Eisenberg, *Science*, **251**, 1481 (1991).
33. J. Deisenhofer and W. Steigemann, *Acta Cryst.* **B31**, 238 (1975).
34. H. Meirovitch and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA*, **78**, 6584 (1981).
35. D. R. Ripoll, L. Piela, M. Vásquez, and H. A. Scheraga, *Prot. Struct. Funct. Genet.*, **10**, 188 (1991).
36. S. Talluri, D. M. Rothwarf, and H. A. Scheraga, *Biochemistry*, **33**, 10437 (1994).
37. A. Wlodawer, L. A. Svensson, L. Sjölin, and G. L. Gilliland, *Biochemistry*, **27**, 2705 (1988).
38. P. Güntert, W. Braun, and K. Wüthrich, *J. Molec. Biol.*, **217**, 517 (1991).
39. G. T. Montelione, K. Wüthrich, A. W. Burgess, E. C. Nice, G. Wagner, K. D. Gibson, and H. A. Scheraga, *Biochemistry*, **31**, 236 (1992).
40. W. A. Hendrickson and M. M. Teeter, *Nature*, **290**, 107 (1981).
41. M. M. Teeter, *Proc. Natl. Acad. Sci. USA*, **81**, 6014 (1984).
42. P. K. Pallaghy, B. M. Duggan, M. W. Pennington, and R. S. Norton, *J. Mol. Biol.*, **234**, 405 (1993).
43. M. Vásquez and H. A. Scheraga, *J. Biomol. Struct. Dyn.*, **5**, 757 (1988).
44. J. Vila, R. L. Williams, M. Vásquez, and H. A. Scheraga, *Prot. Struct. Funct. Genet.*, **10**, 199 (1991).
45. D. M. Rothwarf and H. A. Scheraga, *J. Am. Chem. Soc.*, **113**, 6293 (1991).
46. D. M. Rothwarf, Y.-J. Li, and H. A. Scheraga, *Prot. Sci.* **4**(suppl. 2) (1995), Abstract 237-S, p. 105.